



*Citation for published version:*

Tekin, B & Soruc, A 2016, 'Using Corpus-Assisted Learning Activities to Assist Vocabulary Development in English', *The Turkish Online Journal of Educational Technology*.

*Publication date:*  
2016

*Document Version*  
Peer reviewed version

[Link to publication](#)

**University of Bath**

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Using Corpus-Assisted Learning Activities to Assist Vocabulary Development in English

**Dr. Bilal TEKIN, Freelance Teacher Trainer**

[bilaltekin66@gmail.com](mailto:bilaltekin66@gmail.com)

**Assist. Prof. Dr. Adem SORUÇ**

*Sakarya University, Turkey*

[ademSORUC@sakarya.edu.tr](mailto:ademSORUC@sakarya.edu.tr)

### ABSTRACT

In recent years, the use of corpora has achieved increasing prominence for educational purposes in general, for second language vocabulary development in particular. To this end, the qualitative study reported in this article sought to find out learners' perceptions related to the use of corpus assisted vocabulary learning activities in an EFL classroom. The study was carried out at an international high school in Istanbul, conveniently sampling 26 learners (16 male, 10 female). The learners were exposed to four unknown words for two classroom hours of activities prepared using a corpus program as in the literature (e.g., Thurstun & Candlin, 1998). Upon completion of the activities, data, first of all, came from a persona evaluation checklist. Then, reflection papers were collected, and at the end all learners interviewed. The qualitative data were transcribed and coded. The results showed that learners found corpus-assisted vocabulary learning activities interesting enough, identifying their thoughts with such expressions as "easy", "fun", "innovative", "autonomous", "practical", but "complex". Bearing all these results in mind, a series of implications are given and suggestions made for both language teachers and learners.

### INTRODUCTION

Recent developments in the field of educational technology have led to a renewed interest in corpus-assisted language learning (Johns 1991; O'Keeffe, McCarthy & Carter, 2007; Flowerdew, 2012), which has been used interchangeably for different learning purposes, with a particular reference providing not only a better road for vocabulary development (e.g., Thurstun & Candlin, 1998; Binkai, 2012) or grammar teaching (e.g., Boulton, 2009; Smart, 2014; Lin & Lee, 2015) or writing skills (e.g., Quinn, 2014), but also a "language-learning activity" (e.g., Gavioli 1997, p. 84), or a self-correcting tool (e.g., Gaskell & Cobb 2004; Yoon 2008; Quinn, 2014). Given that corpus assisted language learning is made up of using technological tools, particularly concordancers, for pedagogical purposes (Gilquin & Granger, 2010), this study set out to explore, rather than measuring the effectiveness of corpus-assisted vocabulary learning activities, learners' perceptions about how they found vocabulary activities prepared using concordancing lines.

### CORPUS-ASSISTED LANGUAGE LEARNING

Corpus with its simplest definition is the collection of written or spoken texts stored on a computer for text analysis (O'Keeffe et al., 2007), and having both direct and indirect applications (Römer, 2011). It is on the one hand directly used as a language learning activity both in and out of classroom; on the other hand indirectly used as a language teaching resource.

The direct use of corpora is most commonly in line with data-driven learning (DDL), a term derived by Tim Johns (1991). DDL is described as a tool by which learners can study a language, simply getting exposed to the input in concordancing lines (Frankenberg-Garcia, 2014), and analyzing 'masses of input in a quest to become more familiar with structural regularities via inductive means' (Geluso & Yamaguchi, 2014, p. 227). DDL can be viewed as a pivotal pedagogical tool on the grounds that it includes large amounts of text, and as it 'involves skills such as scanning and comprehending and observing recurring patterns, and forming conclusions based on these reoccurrences' (Quinn, 2014, p.167). In other words, learners explore data by themselves detecting 'patterns among multiple language samples' (Boulton, 2010, p. 535). Put simply, learners take advantage of corpora which allows them to discover either alone or in

pairs driving them both to notice and to identify linguistic features in data (Lin & Lee, 2015) – which is closely related to Schmidt's (1990) noticing hypothesis as learners focus on either the form or the word itself in the concordancing lines.

Corpus-assisted language learning is known to have various advantages such as facilitating multi-contextual lexical acquisition (O'Keeffe, McCarty, & Carter, 2007), bringing authenticity to classroom (Gilquin & Granger, 2010), and contributing to autonomous learning (Binkai, 2012). On the other hand, it is also known that concordancing lines lined up vertically with the words on their right and left side startle learners because of their physical appearance (Lamy & Klarskov Mortensen, 2012), challenging learners to interpret the concordance lines (Koosha & Jafarpour, 2006), and taking so much time processing the meaning of the word (Yoon, 2011; Quinn, 2014; Lin & Lee, 2015). Corpus-assisted language learning is however a well-known tool or a 'promising technique' (Gilquin & Granger, 2010, p. 367) by which learners contact with authentic language and discover either linguistic forms or lexis and develop important cognitive skills (Gilquin & Granger, 2010).

The role of corpus-based instruction on vocabulary development has been investigated to date. To take a specific example, Thurston and Candlin (1998) introduced concordancing program, *Microconcord*, which contains 1.016.000 words taken from academic texts to teach learners frequently used academic English words. In their study, examples were, firstly, presented to the learners in concordance lines, which were then followed by a series of tasks/activities to ensure that the learners used the targeted items correctly and appropriately. Though learners were initially startled by the cut-off sentences of one-line concordances, both learners and teachers later admitted that such tasks based on corpora helped them to gain 'helpful, very different and innovative' (p. 277) insight into vocabulary development. Thurston and Candlin (1998) therefore argued that the rich language context obtained by corpora, especially when the words are studied or exercised, gives learners 'considerable opportunities' (p. 278) to broaden both their lexical knowledge and/or grammatical awareness.

Koosha and Jafarpour (2006) investigated the role of DDL and concordancing materials in the production of collocation of prepositions by English-major Iranian learners (n=200) in three different universities. The participants were divided into two groups, and related to prepositions and their collocational patterns while one received traditional instruction, the other, DDL-based instruction. The results showed that learners receiving DDL instruction outperformed those in the traditional group, which thereby further showed that DDL had a greater "explanatory power" (p. 202) compared to traditional ways that teachers generally use in the classroom.

In another study, Binkai (2012) conducted a study on vocabulary learning using concordancers of a corpus (New College English Corpus and BNC). Then, to measure how much learners enjoyed corpus-driven vocabulary learning, Binkai gave a questionnaire to Chinese learners (n=87). The results showed that a large majority of learners either enjoyed (66.7%) or liked (29.8%) learning vocabulary using the corpus. When compared to traditional ways, this time many learners found corpus-driven way helpful (54%), although some (24.2%) argued both ways are similar. In addition, although some learners sometimes "felt dull and lost the interest" (p. 135), a large number (94.2%) still reported that their vocabulary knowledge improved to a greater extent, as 'concordancing lines are of great help in discovering the collocation, colligation and prosody of the search word' (p. 135).

More recently, Quinn (2014) investigated how third-year tertiary learners (n=58) in an intermediate-level EFL writing course used corpora for the purpose of self-correcting teacher coded errors. After five-class introduction to general corpus usage, the learners revised texts familiarizing themselves with the process. The questionnaire data showed that the use of corpus as a learning activity helped learners to 'write more naturally' (p.173).

Yılmaz and Soruç (2015) researched the effectiveness of a corpus program (COCA) on vocabulary development, assigning some learners (n=20) into the corpus group, others (n=20) into the control group. In an-eight-classroom hours of instruction, to learn 23 target unknown words (measured by a word recognition test prior to instructions), while learners in the corpus group were exposed to concordance lines, those in the control were engaged in traditional vocabulary activities such as finding synonyms, antonyms, etc. The word recognition test showed that the corpus group

scored significantly more than the control group; interviews that learners enjoyed learning by the corpus program as it provided them words in “repetitive context” (p. 2630).

Several studies have showed that corpus-based programs and/or DDL is effective on the development of grammatical features (e.g., Boulton, 2009; Smart, 2014; Lin & Lee, 2015), lexis (Thurstun & Candlin, 1998; Binkai, 2012), writing skills (e.g., Quinn, 2014), collocation analysis (Frankenberg-Garcia, 2012), and error-correction (Gaskell & Cobb 2004). But it is still rare to find data about what corpus studies call for the language classroom pedagogy and what teachers do in the classroom (Mukherjee, 2006). More important, a large body of research remained limited to tertiary education (Chambers, 2007; Braun, 2007). This present study therefore seeks to find, but not limited to, the answers of the following research questions:

RQ1: What do high school learners think about corpus-assisted vocabulary learning activities?

RQ2: How much do learners feel satisfied with corpus-assisted vocabulary learning activities?

## THE STUDY

### Setting and Participants

This study was carried out at an international high school in Istanbul at the beginning of fall term in 2015-2016 academic years. At the school 350 learners from approximately twenty nationalities were studying from kindergarten to Year 12. The school was a high-tech school, having technologically equipped classes as well as two computer laboratories. The learners were taking more than six hours of English classes, all taught by native American teachers. It was Year-9 learners (N=30) with ages ranging from 13 to 14 (from 11 nationalities) that were conveniently chosen for the study. Prior to the study, the main research goal was explained, and confidentiality was assured. Then, learners were asked to sign that they agreed to participate to the study used for research purposes and possible publication. Except for two students, all were happy to sign and to participate. And as another two missed the instructions, it left 26 learners, of whom sixteen were males (61.5%); ten were females (39.5%).

### Selection of the Target Words

The school curriculum followed Cambridge International Education (CIE), and the school chose an upper-intermediate English course book for learners at the level. To determine the target words, Michigan Proficiency Vocabulary Test (ECPE C2) with 50 vocabulary items was given, according to which words known at and above 7,7% were excluded, leaving at the end 4 target words (see figure 1 as a sample) to be taught by using corpus-assisted vocabulary learning activities (see sample activities below).

words	Correct answer	%	words	Correct answer	%
proponent*	0	0.0	painstakingly	2	7.7
dispel*	0	0.0	voraciously	2	7.7
denigrate*	1	3.8	curb	3	11.5
vie*	1	3.8	adept	3	11.5
crease	2	7.7	looming	5	19.2

Figure 1. The selection of the words (with asterisk) below 7, 7 per cent.

### Instructional Treatment

After the target words were selected, learners were initially trained on how they could take advantage of corpus-assisted vocabulary learning activities. To illustrate, learners were first introduced to British National Corpus – a free-to-access online database containing 100,000,000 English words, and then given some information about how they could search and use themselves concordance lines to learn the meaning of unknown words.

After training was completed, learners were handed out vocabulary activities prepared using concordance lines and following those in Thurston and Candlin (1998) – all prepared to expose learners to the targeted four words in two intact regular classroom hours. Throughout the instructional period, learners received same number and type of activities for all four of the words. They were, for instance, asked to have a look at the concordances and to “familiarize” themselves and to practice at the end (see sample activity types for the word “vie” below).

#### Activity 1. Concordance lines to guess the meaning of the word “vie”

LOOK		
Read the sentences below and try to infer the meaning of the target word.		
carefree surface bitter enmities stir. Factions at the Evercourt	vie	for the favour of the Queen. Old rivalries are barely submerged
let her out canvassing?) They ask these wives, who dutifully	vie	with one another in supportive dullness, many questions, and
rents as an excuse to throw a party, during which they would	vie	with each other in largesse, showing off their wealth as well as
re to fix the eye. Symbols and echoes of celebrated canvases	vie	with original designs and lush flourishes to create an aesthetically
aurants. Spanish, Chinese, French, Indian, and McDonalds all	vie	for attention against English-run establishments. For those de

#### Activity 2. Concordance lines to familiarize learners with the word “vie”

FAMILIARIZE						
Read the concordance lines again and decide which of the following sentences are true.						
To vie	involves	competing	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	hoping	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	contending	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	striving	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	sleeping	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	criticizing	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	supporting	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	struggling	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>
To vie	involves	challenging	TRUE	<input type="checkbox"/>	FALSE	<input type="checkbox"/>

After learners were familiar with the meaning of the targeted word, they were given further activities to consolidate the use of the word, for instance, “vie”. Therefore, they were strongly advised to look at the concordance lines, and encouraged to be more careful about the language surrounding the key word. To make them practice, activity 3 was given, and asked learners to decide whether “vie” goes with “for” or “with” by getting help from the activity 1.

PRACTICE	
<b>Find the missing word. Use either “vie for” or “vie with”.</b>	
me apparent whether the case of Ian X was unique. # Landlords	vote of Tweeddale tenants # By William Chisholm # SCO
in character'. It expected that Scott would produce a building to	his great designs abroad and concluded with the hope that'
oint an unrepresentative number of Tories for Scotland. # Rivals	control of home rule campaign # By Joanne Robertson and
# Rosa' Ballerina', pink sidalcea, blue catananche and begonias	attention Along the hedge that backs onto the lane is a mag
pronouns -- or rather, they are created by the pronouns. Critics	each other in discovering the dominant qualities of the Son
concluded.' Given the acrimony between the parents who both	the children's affection and blacken each other' s character
from the full body to three lone penises (gilded). These examples	other reminders of some of the more arcane uses to which

*Activity 3.* Concordance lines to practice prepositional uses of the word “vie”

After learners were familiarized with filling gaps, or confirming the meaning of the words doing true/false activities, to increase their familiarity further as in Thurstun and Candlin (1998), matching exercises were given as well (see activity 4).

*Activity 4.* Concordance lines to match parts of sentences using the word “vie”

PRACTICE	
<b>Draw lines matching parts of the sentence on the left hand side with the appropriate part on the right using “vie for” or “vie with”.</b>	
g to convince herself that she is still young and nubile enough to	a strategic yet friendly investment in Jaguar. That would co
ated. Their first attempts to invent an anti-hero strong enough to	each other to advertise their copious offerings of wine. To
ly sloshed, and you stay that way longer. No wonder the airlines	the title Miss Golden Gate. In last summer's run, actress L
es is ambitious. A more likely scenario is that GM and Ford will	the popularity of Tarzan and Superman, and soft enough t

In the final group activities, learners were asked to write sentences using the target word such as “vie”, generally either creating a new sentence (activity 5) or paraphrasing the sentence (activity 6), so that they could produce a sentence using and testing their knowledge of the target word as well as noticing their different forms and meanings thanks to the earlier activities.

*Activity 5.* Supplementary words to produce a new sentence using the word “vie”

Read the following expressions and make up sentences using “vie for” or “vie with”.				
Samsung	mobile phone	Apple	10 years	best
.....				
Mercedes	BMW	market	car	comfortable
.....				

Activity 6. Rewrite the sentences using the word “vie”

**CREATE**

Use the words you have learnt in this exercise. Read the following sentences and rewrite them using “vie for” or “vie with”.

GPs were encouraged to compete for more patients by offering higher standards of care

It was encouraged that .....

A welding machine designed and produced by Jua Kalis is beginning to compete with the conventional, imported Japanese product

.....

With these series of activities as suggested by Thurstun and Candlin (1998), instructional period was completed, pushing learners into being engaged in corpus-assisted vocabulary activities based on the following processes as “look”, “familiarize”, “practice”, and “create”.

**Data Collection Instruments and Procedure**

As soon as the instruction was over, to explore learners’ perceptions related to corpus-assisted vocabulary learning activities, three data collection instruments were used: reflection papers, semi-structured interviews and persona evaluation scale – all were to triangulate “the validity and reliability of the information” (Johnson, 1992, p. 146). In the reflection paper, learners were asked twice immediately after classes to write what they thought about corpus-assisted vocabulary learning activities, how they found concordances to learn vocabulary, and whether the activities sufficed to develop their vocabulary knowledge. They all wrote.

In addition, a persona evaluation five-point Likert scale (Fig. 2) was developed to explore how much learners felt satisfied with doing the activities, changing their responses from the least satisfaction level (1= Not at all satisfied) to the most satisfaction level (5= Extremely satisfied).

1	2	3	4	5
not at all satisfied	slightly satisfied	moderately satisfied	very satisfied	extremely satisfied

Fig. 2. A persona evaluation scale to measure learners’ level of satisfaction

Finally, to collect further qualitative support, 10 learners were randomly selected for semi-structured interview sessions – all were conducted immediately after classes meeting with learners individually in an office at the school and showing learners the activities done in the classroom to remind them. It was semi-structured interview, because it was thought for the study to be ‘less rigid’ (Mackey & Gass, 2005, p. 173) and the interviewer (the first author) could use a series of written questions to guide him on the one hand, and could feel free to probe into new topics. Some of the main questions, however, addressed in the interviews were as follows:

1. How effective did you find corpus-assisted vocabulary learning?
2. What did you like and dislike most?
3. Did you learn or not? Why?
4. How did you feel when you were doing the exercises?



### Data Analysis

To analyze the qualitative data, a grounded approach (e.g., Glaser & Strauss, 1967; Dörnyei, 2007) was adopted for the analysis of learners' comments collected from reflection papers and of those transcribed from the interviews. According to this approach, the collected data was first analyzed by open coding, an 'interpretive process by which data are broken down analytically' (Corbin & Strauss, 1990, p. 12), followed by axial coding, where the researchers associated the categories with their subcategories and analyzed the relationships among categories, and lastly, by selective coding, where one or more core categories were selected to generate main theme that linked the categories. In the analysis stage, between the researchers some discussions were made for whether to add one more theme to the found already (see the themes found below); but thanks to "peer debriefing" sessions occurring 'between the evaluator and a "disinterested peer" (*author's quotations*) concerning the evaluation findings, conclusions, and tentative hypotheses' (Lynch, 1996, p. 67), the disparities about the inclusion of a possible theme were negotiated and resolved. As to the analysis of persona evaluation scale, descriptive statistics were run; namely, percentages of learners' responses were calculated.

### RESULTS

#### Results – Interview and Reflection Data

According to the analyses, five common themes were explored, all describing corpus-assisted vocabulary learning activities: Innovative, autonomous, easy and fun, practical, and complex.

#### Innovative

It was extracted from both reflection papers and semi-structured interviews that learners had their best interests at heart about corpus-assisted vocabulary activities, notably pointing out that they had never used before such a learning tool as corpus to assist them to develop their vocabulary knowledge in English.

Many learners (n=15) stated that they found corpus and concordance a new and a useful vocabulary learning equipment. Some, for instance, specifically stated that

"Concordance helped me know how to expand my knowledge and vocabulary in ways I didn't know about."

"It was very effective and new because it changed the way I learn words. It was very helpful."

"It is really good if you need a sample sentence to help define a word. It is original."

"Concordance will be very useful in the 21<sup>st</sup> century when people need to learn new things from computer. And this tool is very new to me."

#### Autonomous

Concordance lines, searching for words, guessing meanings, confirming/disconfirming the hypotheses – all helped learners to work either individually or in pairs, and to feel more responsible and more autonomous and more independent about their learning, which might not have been attained easily by doing traditional vocabulary learning activities. Learners' responses (n=17) showed that corpus-based activities enhance autonomy, some of which are given as follows:

"This method is probably the best way of learning because it is very understandable and can be done anywhere at home for instance by yourself."

"From now on it is my favorite method of learning a new word. I will use this method next time I learn a new word."



“I found out that there is a new website (BNC) where I can search and learn new words by myself.”

“It helped me to gain great personal initiative. I started learning by myself.”

“I think upper intermediate learners should take charge of their own learning; therefore, I liked this type of activities.”

### **Easy and fun**

The learners also had a lot of fun doing corpus-assisted vocabulary activities. For instance, a large majority of learners (n=24) expressed that they enjoyed concordance lines and related activities. However, similar number of learners still argued that without training, the concordance tool and/or activities would be meaningless and dull. Some specific expressions were as follows:

“I had a great fun doing the activities, searching in the lines and trying to understand the sentences.”

“It is really hard to understand unless you have a tutor. Without anyone to help you, it is very hard to use it.”

“I felt smart while I was doing the exercises. I learned a new word in a few minutes without struggling because it is easier to use rather than a dictionary where you have to look through all those pages.”

“I felt weird since I didn’t understand what they were about but I later understood what they were about and I enjoyed it.

“I felt great and energized as if I drank two large cups of coffee.”

### **Practical**

Learners grasped not only the meaning, but also use of the words. Specifically, they learned collocations; that is, knowledge of how words are essentially combined to build blocks of words. Some learners (n=12) similarly stated, for instance, that

“I liked concordance lines because it helped me improve my English, especially my knowledge of words going together in the sentence. It was simple”

“It gave me great opportunity to practice not only the meaning of new words but also the structure of sentences, tenses, and sentence context.”

“It was an efficient or practical way to use these daily new words in the activities rather than using traditional ways to know their definition.”

“Doing corpus-based activities was very workable for me. They were doable.”

### **Complex**

Although a large number of learners favored corpus-assisted learning activities and liked learning vocabulary doing these activities, almost all learners (n=21) still argued that concordance lines, generally in abbreviated or in an uncompleted form, made it difficult for them to understand the real meaning of the sentence, given further the fact that the sentences in the corpus were collected spoken and written data from native speakers of English. In other words, they found these

halved sentences confusing. It was confusing, because learners could not bridge the meaning lying behind the word targeted and that in the front, which led many learners to spend much more time than expected. Some statements are as:

“Many distinct sentences are shown, which overstressed me.”

“It is really hard to find the one you are looking for in the sentences.”

“Sentences were a little bit difficult for me to interpret.”

“I can’t combine the meaning in the sentences.”

“These sentences or words are not for me, they are for native speakers. I paid great effort to understand them.”

### Results – Persona evaluation scale

Persona evaluation scale was additionally given to explore how much the learners were satisfied with corpus-assisted vocabulary activities. Only percentages of responses were calculated to find out learners’ satisfaction level. According to the results, it revealed that with corpus-assisted vocabulary learning activities

- 69.2 % of the learners (n=18) were extremely satisfied
- 19.2 % (n=5) very satisfied
- 11.6% (n=3) moderately satisfied

Although a large number of learners (n=21) both wrote in their reflection papers and stated in the interviews that concordance lines caused remarkable difficulty or complexity for them to grasp words or the meaning of the whole sentence, they nevertheless did not show any level of dissatisfaction in the persona evaluation scale. That is, all learners were found satisfied enough with the activities prepared using corpus although the activities were still argued to be difficult, or sometimes complex. To conclude, the results of the satisfaction scale corroborated those of interviews, and reflections of learners.

### DISCUSSION

As can easily be noticed from the design of the study, this research set out to explore learners’ perceptions of and attitudes to corpus-assisted vocabulary learning activities, not their effectiveness on vocabulary development in second language,. Data gathered by reflection papers, interviews, and satisfaction scale showed us that learners favored corpus-based activities when learning or discovering the meanings of new words to them, describing the activities with the words as “innovative”, “easy and fun”, “practical”, and “explorative”. Namely, not only did they state that they liked doing the activities, but claimed that they wanted to use corpus in the future when learning new words.

Being an “innovative approach” (Thurstun & Candlin, 1998, p. 277) to second language vocabulary development and providing “a rich experience” (ibid.: p. 277) with multiple examples of words, word groups or collocations in authentic examples, corpus-assisted vocabulary activities helped learners in this study both discover the meaning of the new words and notice different uses of targeted words in different sentences. Schmidt (1990) argued that ‘noticing is the necessary and sufficient condition for converting input to intake’ (1990, p. 129), which might be done only ‘when the demands of a task focus attention on what is to be learned’ (ibid., p. 129). As many learners (n=17), for instance, pointed out that they felt more responsible about their own learning because the activities encouraged them to do further research, they noticed many more different uses of targeted words in the sentences, which may lend support to Schmidt’s (1994) noticing hypothesis.

Authenticity is also an important matter of fact that teachers should consider well, which, according to the results of this study, was achieved by corpus-assisted vocabulary learning activities as with the results of the earlier research (e.g., Gilquin & Granger, 2010; Yoon, 2011). Authenticity appeared, because the more learners were engaged in doing the activities, the more they were exposed to the real uses of the targeted words in authentic sentences either written or spoken by native speakers of English, thus helping learners to notice and practice words in their real uses and in their different forms (e.g., Frankenberg-Garcia, 2014; Leel, 2011). This finding was encountered in the interviews made with the learners, arguing for instance that *“it helped me greatly broadening my vocabulary knowledge in a correct and original way”*, *“it enabled me to produce sentences from the original sentences”* and *“I learned how native speakers used the words in sentences”*. Teachers should therefore think about the degree of authenticity of the materials that they take into their classroom, and seek for much better ways to operationalize authenticity, or to incorporate corpus into material development.

In addition, corpus-assisted learning activities were found to enhance autonomy of the learners, as they found many opportunities ‘to control their own learning’ (Binkai, 2012, p. 135) or to study in pairs working out the meanings of the words. As learners were more engaged, they were more motivated to learning words using a new tool, and by doing so with higher motivation learners gained an ability to become autonomous or an ability to exercise autonomy (e.g., Cotterall, 2008; Wenden, 1991), which Holec (1981) defined as “the ability to take charge of one’s own learning” (p. 3). Given that in today’s classroom as learners generally have smart phones, teachers could take advantage of mobile learning preparing corpus-assisted activities; learners, therefore, may not be restricted to school or classroom only. For instance, one of the learners stated similarly in the interviews that *“I am definitely going to use this tool at home while learning new words. This refreshed my mind on learning new words and collocations.”*

Despite having reached its results by qualitative data collection instruments, the present study can still get support from the similar results of the quantitative studies in the literature. For instance, Geluso and Yamaguchi (2014) asked Japanese learners (n=30) at tertiary level to search for words and phrases using a corpus tool named as Corpus of Contemporary American English (COCA), and their results from a questionnaire, interviews, and reflection blogs showed that although having some “reservations” (p. 240), the learners still stated that they liked using corpora, and described it as “a good tool” (p. 240). Yılmaz and Soruç (2015) likewise reached similar conclusions about the effectiveness of using corpora for vocabulary development, finding out that learners exposed to corpus-based instruction performed much better on the word recognition test than those receiving traditional instruction. In another study, Çelik (2011) used two instructional types, namely comparing corpus to online dictionary use, involving 68 learners at tertiary level, to teach ten collocations in five sessions. Pre/posttest results showed that learners receiving vocabulary instruction through corpora remembered much more than those engaged in online dictionary use.

The use of corpora has been confirmed not only in teaching vocabulary (e.g., Thurston & Candlin, 1998; Çelik & Keser, 2010) or grammar (e.g., Lin & Lee, 2015), but in developing language skills as well. For instance, Yoon (2008) defined corpus a tool serving as a “meaningful reference” and/or as a “catalyst” (p. 44) to mediate learners in their writing, and to enhance learners’ awareness especially related to “common usage and collocation in writing.” (p. 44).

Nevertheless, such type of corpus activities was not without problems. One potential weakness was that learners are more likely to be confused by physical appearance of concordancing lines (Lamy & Klarskov Mortensen, 2007). At first sight, although the learners in this study were trained about concordances, they could not develop immediate strategies to further their vocabulary knowledge. Nor could they extract instantly the real meaning of the targeted word, either because of the complexity of the words or because of the halved sentences. Therefore,

- teachers should bring specifically prepared and carefully selected sentences into the classroom.
- complexity, difficulty, and abstractness of the words should be well thought when preparing corpus-assisted vocabulary activities.

- corpus-based activities should not be the single source of activity; they should be, rather, given whether to provide a fresh start or to meet learners' needs.
- strategy training should be given to learners as to how they can manage to establish the meaning of the sentence or to guess the meaning of the word.

The role of emotions and psychological factors in language learning, or affective aspects, are of greater importance in second language classroom (e.g., Brewer, 2010, 2013; Dewaele, 2011). Individual learner differences such as motivation, learner beliefs, attitudes, personality features, and identities (e.g., Soruç & Griffiths, 2015, Dörnyei & Ushioda, 2009) – all these differences can be related to affect, or emotional level of learners. This situation was also found in the responses of learners gathered by persona evaluation or satisfaction scale, according to which *all* learners were satisfied with the corpus-based activities. Surprisingly never did a learner state that s/he was not satisfied with the activities. It was due to the fact that the learners felt satisfied, they approached to the activities in a positive manner.

Even though this study was conducted in an EFL setting, the following implications can still be useful for both EFL and ESL learners or teachers.

- Language teachers should find ways to enhance learners' autonomy to make learners use their own potential.
- Excessive or repetitive usage of concordance lines can make learners bored; the instruction monotonous (Thurstun & Candlin, 1998). Therefore, learners' needs should be thought carefully.
- Different ways to mobile learning could be sought, as it can be found simply in today's classrooms.
- Gilquin & Granger (2010) claims the reason why teachers do not benefit from corpus assisted language learning is that they do not know how to take advantage of corpora. Teachers could be trained as to how they can use corpus effectively in the classroom, so that they can have corpus pedagogical knowledge.
- Proficiency level of learners should be carefully considered when preparing or using corpus-assisted vocabulary learning activities. It would be much more efficient in the upper-intermediate and advanced levels.

## Conclusion

Although this study is qualitative, having aimed to explore learners' qualitative perceptions of and attitudes to corpus-assisted vocabulary learning activities, and its results are limited to the data collected from one setting in a short instructional period of time, its findings are still consistent with those in the literature (e.g., Thurstun & Candlin, 1998). All learners were found satisfied with the activities; a large majority found the activities innovative, easy and fun, and practical. Although learners were puzzled by the concordancing lines at first sight, they still liked doing the activities on their own. Future research studies can still be carried out to further investigate, if any, greater effectiveness of corpus-assisted learning activities on vocabulary improvement designing more controlled experimental groups.

## REFERENCES

- Binkai, J. (2012). An Empirical Study on Corpus-driven English Vocabulary Learning in China. *English Language Teaching*, 5(4), 131-137.
- Black, P. (1999). Assessment, learning theories and testing systems. In P. Murphy (Ed.), *Learners, Learning and Assessment*. London: Open University Press.
- Boulton, A. (2009). Testing the limits of data-driven learning: Language proficiency and training. *ReCALL*, 21(01), 37.
- Boulton, A. (2010). Data-Driven Learning: Taking the Computer Out of the Equation. *Language Learning*, 60(3), 534-572.

- Braun, S. (2007). Integrating corpus work into secondary education: From data-driven learning to needs-driven corpora. *RECALL*, 19(3) 307-328.
- Chambers, A. (2007) Popularising corpus consultation by language learners and teachers. In: Hidalgo, E., Quereda, L. and Santana, J. (eds.), *Corpora in the foreign language classroom*. Amsterdam: Rodopi, 3–16.
- Corbin, J. M., & Strauss, A. (1990). Grounded theory research: Procedures, canons, and evaluative criteria. *Qualitative sociology*, 13(1), 3-21.
- Dörnyei, Z. (2007). Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies. Oxford University Press.
- Flowerdew, L. (2012). *Corpora and language education*. New York: Palgrave Macmillan.
- Frankenberg-Garcia, A. (2012). Learners' use of corpus examples. *International Journal of Lexicography*, 25(3), 273-296.
- Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, 26(2), 128-146.
- Gaskell, D., & Cobb, T. (2004). Can learners use concordance feedback for writing errors? *System*, 32(3), 301-319.
- Gavioli, L. (1997). Exploring texts through the concordancer: Guiding the learner. In A. Wichmann, S. Fligelstone, T. McEnergy, and G. Knowles (Eds.), *Teaching and Language Corpora* (pp. 83–99). London: Longman.
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Learner attitudes and efficacy. *ReCALL*, 26(02), 225-242.
- Gilquin, G. and S. Granger. (2010). 'How can data-driven learning be used in language teaching?' in A. O'Keeffe and M. McCarthy (eds.). *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- Glaser, B. S., & Strauss, A. (1967). *The discovery of grounded theory*. New York.
- Johns, T. F. (1991). Should you be persuaded: two samples of data-driven learning materials. *ELR Journal*, 4, 1-16.
- Johnson, D. M. (1992). *Approaches to research in second language learning*. New York: Longman.
- Koosha, M. and Jafarpour, A. (2006). Data-driven learning and teaching collocation of prepositions: the case of Iranian EFL adult learners. *Asian EFL journal quarterly*, 8(4): 192–209.
- Lamy, M.N., & Klarskov Mortensen, H. J. (2012). Using concordance programs in the modern foreign languages classroom. Module 2.4. In Davies, G. (ed.), *Information and communications technology for language teachers (ICT4LT)*. Slough: Thames Valley University. [http://www.ict4lt.org/en/en\\_mod2-4.htm](http://www.ict4lt.org/en/en_mod2-4.htm)
- Leel, H. (2011). In Defense of Concordancing: An Application of Data-Driven Learning in Taiwan. *Procedia - Social and Behavioral Sciences*, 12, 399-408.
- Lin, M. H., & Lee, J. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, 69(3), 264-274.
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge: Cambridge University Press.

- Mackey, A., & Gass, S. M. (2005). *Second Language Research: Methodology and Design*. Lawrence Erlbaum Associates: New Jersey.
- Mukherjee, J. (2006). Corpus linguistics and language pedagogy: The state of the art – and beyond. In S. Braun, J. Mukherjee & K. Kohn (Eds.), *Corpus technology and language pedagogy New resources, new tools, new methods* (pp. 5–24). Frankfurt: P. Lang.
- O'Keeffe, A., McCarthy, M. & Carter, C. (2007). *From corpus to classroom: Language use and language teaching*. Cambridge: Cambridge University Press.
- Quinn, C. (2014). Training L2 writers to reference corpora as a self-correction tool. *ELT Journal*, 69(2), 165-177.
- Römer, U. (2010). Using general and specialized corpora in English language teaching: Past, present and future. In M.C. Campoy-Cubillo, B. Bellés-Fortuño, & M.L. Gea-Valor, (Eds.), *Corpus-based approaches to English language teaching* (pp. 18-38). London: Continuum International Publishing Group.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129-58.
- Schmidt, R. W. (1994). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165-210). London: Academic Press.
- Smart, J. (2014). The role of guided induction in paper-based data-driven learning. *ReCALL*, 26(02), 184-201.
- Soruç, A., & Griffiths, C. (2015). Identity and the spoken grammar dilemma. *System*, 50, 32-42.
- Thurstun, J., & Candlin, C. N. (1998). Concordancing and the teaching of the vocabulary of academic English. *English for Specific Purposes*, 17, 267-280.
- Yoon, H. (2008). More than a linguistic reference: The influence of corpus technology on L2 academic writing. *Language Learning and Technology*, 12(2), 31-48.
- Yoon, C. (2011). Concordancing in L2 writing class: an overview of research and issues. *Journal of English for Academic Purposes*, 10(3), 130-139.
- Cotterall, S. (2008). Autonomy and good language learners. In C. Griffiths (Ed.), *Lessons from good language learners* (pp. 110-120). Cambridge: Cambridge University Press.
- Wenden, A. (1991). *Learner strategies for learner autonomy*. Hemel Hempstead: Prentice Hall.
- Geluso, J., & Yamaguchi, A. (2014). Discovering formulaic language through data-driven learning: Student attitudes and efficacy. *ReCALL*, 26(02), 225-242.
- Çelik, S. (2011). Developing collocational competence through web based concordance activities. *Novitas-ROYAL (Research on Youth and Language)*, 5(2), 273-286.
- Çelik, S., & Keser, H. (2010). The Correlation between Learners' Logs of Navigations through Online Corpora and Lexical Competence. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, 43(2), 149-170.
- Lin, M. H., & Lee, J. (2015). Data-driven learning: Changing the teaching of grammar in EFL classes. *ELT Journal*, 69(3), 264-274.
- Yılmaz, E., & Soruç, A. (2015). The use of Concordance for teaching Vocabulary: A data-driven learning approach, *Procedia Social and Behavioral Sciences*, 191, 2626 – 2630.